Yan Yan

1005387582

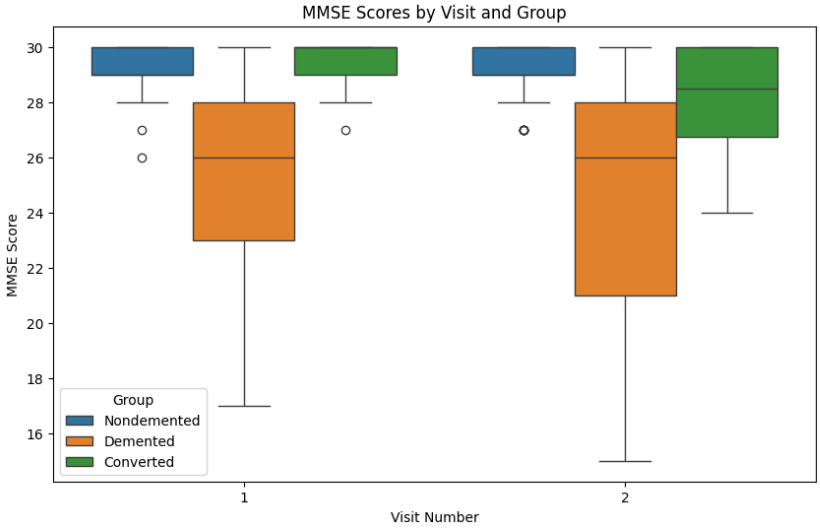
Research question:

How does cognitive function, as measured by the MMSE, change over time within individuals, and how do these changes differ between individuals with and without dementia?

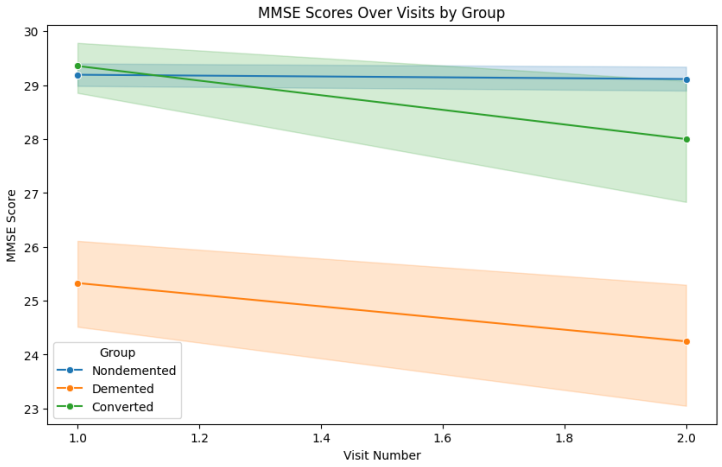
EDA:

The boxplot compares MMSE scores between three different groups (Nondemented, Demented, Converted) over two visits. MMSE scores are a standard measure used to assess cognitive function, with higher scores indicating better cognitive function. The Nondemented group has high median MMSE scores in both visits, with little change between visits. The scores are tightly clustered, indicated by the shorter box, especially in the second visit, which suggests less variability in MMSE scores among this group over time. The Demented group shows noticeably lower median MMSE scores compared to the Nondemented group in both visits, which is expected as this group represents individuals diagnosed with dementia. There's a slight overlap in the interquartile ranges between the first and second visits, implying a stable or slight decrease in scores. The Converted group appears to have MMSE scores that are intermediate between the Nondemented and Demented groups. The term Converted could imply these individuals transitioned from a diagnosis of Nondemented to Demented between visits. There's a noticeable drop in the median score from the first to the second visit, which might indicate cognitive decline over time. There are outliers in both visits for the Nondemented and Converted groups. These outliers are below the main cluster of data points, indicating individuals with significantly lower MMSE scores than the group median. The line plot depicts MMSE score trends across two visits for three groups—Nondemented, Demented, and Converted—highlighting the stability of cognitive function in the Nondemented group, a slight decline in the Demented group the Converted group, potentially indicating a transition in cognitive status. The plot suggests consistent cognitive performance among the Nondemented individuals, greater variability, and a downward trend for the Demented, and a distinct shift for the Converted group that aligns them closer to Demented scores over time, offering valuable insights for the research on cognitive function changes associated with dementia. The provided data table below presents MMSE score statistics, revealing that across two visits, all groups saw a decline in average scores, Initially, both the Converted and Nondemented groups showed high cognitive function with tightly clustered scores,

while the Demented group displayed significantly lower cognitive function and greater variability. By the second visit, the Nondemented group maintained relatively stable and high cognitive function, underscoring distinct cognitive trajectories among the different groups in the context of aging and dementia.



Visit	Group	mean	std
1	Converted	29.36	0.93
	Demented	25.33	3.32
	Nondemented	29.19	0.85
2	Converted	28	2.09
	Demented	24.25	4.4
	Nondemented	29.11	0.96



ANOVA interpretation

The Group (between-subjects) factor has a substantial effect on MMSE scores, with a sum of squares (SS) of 1328.421, mean square (MS) of 664.211, and a very high F-statistic of 56.212. The p-value is reported smaller than 0.01, which indicates that the effect of group on MMSE scores is statistically significant. The Visit (within subjects) factor also influences MMSE scores, with an SS of 22.378, an MS of 22.378, and an F-statistic of 8.859. The p-value for the Visit factor is 0.003, indicating a significant effect of time (Visit) on MMSE scores. The interaction between Group and Visit has an SS of 17.000, an MS of 8.500, and an F-statistic of 3.365. The interaction effect has a p-value of 0.037, suggesting that the way MMSE scores change over visits differs significantly between groups. The ANOVA results indicate that both the group status and the time of visit significantly impact MMSE scores, and there is a significant interaction between these two factors. This means that the changes in cognitive function over time as measured by the MMSE are not uniform across groups. Instead, they vary depending on whether individuals are in the Nondemented, Demented, or Converted group, aligning with the research question that seeks to explore cognitive changes over time and across different dementia statuses.

ANOVA Summary			
Source	**SS**	**MS**	**F**
Group	1328.42	664.21	56.21
Visit	22.38	22.38	8.86
Interaction	17	8.5	3.37

The Shapiro-Wilk statistic for visit 1 is approximately 0.787, with a p-value of about 1.76e-13. And for visit 2 is approximately 0.761, with a p-value of about 5.43e-14. In both cases, the p-values are well below the common alpha level of 0.05, leading to the conclusion that the data does not follow a normal distribution (indicated by the "False" in the 'normal' column). This non-normality of MMSE score distribution at both visits could be due to various reasons such as the presence of outliers, skewness in the data, or the fact that the MMSE scores are bounded between 0 and 30, which can naturally lead to non-normal distributions especially in smaller sample sizes or in samples with a high proportion of high or low scores.

Visit	W	P value	normal
1	0.7875	1.76E-13	FALSE
2	0.7607	5.43E-14	FALSE

The power curve provides a visual representation of how statistical power— the probability of detecting a true effect when it exists—escalates as the number of observations in a study increases, particularly for an anticipated effect size of 0.70. In the realm of cognitive function research, where the MMSE scores are used to discern changes across different groups and over time, this curve is instrumental in determining the requisite sample size for robust statistical analysis. The curve's asymptotic approach to a power of 1 indicates that with a sample size nearing 80 participants, the study is highly likely to have sufficient statistical power to detect a large effect size, thereby reducing the risk of Type II error. This is especially pertinent when evaluating the efficacy of interventions or the progression of cognitive impairment in conditions like dementia, where detecting even subtle changes is crucial for timely and accurate conclusions. If researchers aim to capture a true effect of the magnitude indicated by an effect size of 0.70, then a sample size of around 80 is recommended to ensure a high likelihood of their tests being statistically significant, thereby validating the research hypothesis with confidence. The appropriate sample size that I computed in my code is 45. This sample size strikes a balance between the need for statistical accuracy and the practical considerations of conducting research, particularly in the field of dementia where participant recruitment may be challenging. Given the high power and significant effect size anticipated, the number is kept relatively low, yet it's substantial enough to confer confidence in the results, allowing researchers to make definitive conclusions about the efficacy of interventions or the progression of cognitive impairment as measured by MMSE scores.